

AD 738860



TM-4885/000/00

A GUIDE TO THE CONVERSE INPUT CONVENTIONS

15 March 1972

D D C
RPG
MAR 27 1972
RELEASER
C

15

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) System Development Corporation Santa Monica, California		2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
2b. GROUP		
3. REPORT TITLE A Guide to the CONVERSE Input Conventions		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report, March, 15, 1972		
5. AUTHOR(S) (First name, middle initial, last name) Antonio Leal		
6. REPORT DATE 15 March 1972	7a. TOTAL NO. OF PAGES 10	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO. DAHC15-67-C-0149	8b. ORIGINATOR'S REPORT NUMBER(S) TM-4885/000/00	
9a. PROJECT NO. c. ARPA Order #1327, Amendment No. 4 Program Code No. 2D30 and 2P10. d.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
10. DISTRIBUTION STATEMENT Cleared for public release; distribution unlimited.		
11. SUPPLEMENTARY NOTES <i>T</i>	12. SPONSORING MILITARY ACTIVITY	
13. ABSTRACT This guide describes in detail how CONVERSE interprets the characters which are input to it from a remote terminal. It explains the various conventions for word, number, and sentence formation, etc., as well as the formation of special syntactic entities.		

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT

UNCLASSIFIED

Security Classification

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) System Development Corporation Santa Monica, California	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE

A Guide to the CONVERSE Input Conventions

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report, March, 15, 1972

5. AUTHOR(S) (First name, middle initial, last name)

Antonio Leal

6. REPORT DATE 15 March 1972	7a. TOTAL NO. OF PAGES 10	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO. DAHC15-67-C-0149	8c. ORIGINATOR'S REPORT NUMBER(S) TM-4885/000/00	
b. PROJECT NO. ARPA Order #1327, Amendment No. 4 c. Program Code No. 2D30 and 2P10. d.	8d. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	

10. DISTRIBUTION STATEMENT

Cleared for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES Tr. -	12. SPONSORING MILITARY ACTIVITY
--------------------------------------	----------------------------------

13. ABSTRACT

This guide describes in detail how CONVERSE interprets the characters which are input to it from a remote terminal. It explains the various conventions for word, number, and sentence formation, etc., as well as the formation of special syntactic entities.

TECHNICAL MEMORANDUM

(TM Series)

The work reported herein was supported by the Advanced Research Projects Agency of the Department of Defense under Contract DAHC15-67-C-0149, ARPA Order No. 1327, Amendment No. 4, Program Code No. 2D30 and 2P10.

A GUIDE TO THE CONVERSE
INPUT CONVENTIONS

by

Antonio Leal

15 March 1972

SYSTEM
DEVELOPMENT
CORPORATION
2500 COLORADO AVE.
SANTA MONICA
CALIFORNIA
90406

The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U. S. Government.



Distribution of this document is unlimited.

15 March 1972

i
(Page ii blank)

System Development Corporation
TM-4885/000/00

ABSTRACT

This guide describes in detail how CONVERSE interprets the characters which are input to it from a remote terminal. It explains the various conventions for word, number, and sentence formation, etc., as well as the formation of special syntactic entities.

TABLE OF CONTENTS

<u>Section</u>		<u>Page</u>
	ABSTRACT	i
1.	INTRODUCTION	1
2.	WORDS	1
3.	NUMBERS	2
4.	QUOTATIONS	3
5.	ENTRY CONTROL	3
6.	END OF SENTENCE	3
7.	DATES AND TIMES	4
8.	ARITHMETIC	5
9.	PRECEDENCE	5
10.	MINUS VS. HYPHEN	6
11.	HELP	7
12.	CHARACTER SET	8
	ACKNOWLEDGEMENTS	10

15 March 1972

1

System Development Corporation
TM-4885/000/00

1. INTRODUCTION

CONVERSE accepts input through a remote terminal in normal English. There are, however, a number of constraints on word and sentence formation, etc, as well as some features which aid the user in communicating with the computer. This guide is intended for the CONVERSE user who wishes a detailed explanation of those constraints and features. It describes how CONVERSE interprets the characters that are input.

2. WORDS

Words begin with a letter (note exceptions below) and may contain other letters, digits, hyphens, or dots for abbreviations. Two hyphens next to each other will be taken as a dash and end the word. Examples of legal words are:

THE
B17
THREE-WAY
LEFT-TO-RIGHT
PRE-
IBM-360
DR.
U.C.L.A.
MR.-MRS.

Words formed according to the above rules may be preceded by a series of digits that is optionally separated by a hyphen. For example:

4-H
1972MODEL
12A-X
3M

A complex word may be formed out of any legal characters (see CHARACTER SET) by enclosing them in square brackets. For example:

[****]
[HELLO THERE!]

The brackets themselves are considered part of the spelling of the word.

3. NUMBERS

A number may be written either with a decimal point (real) or without a decimal point (integer). It must not exceed the storage capacity of the computer. An integer must be within the range $\pm 2^{31} - 1$ and a real number within $\pm 10^{+75}$. The decimal point may appear before or after a real number. For example:

53 94.8
-4 .948
0 948.

Only 8 places of precision are accepted in a real number. If more than 8 significant figures appear in the number, only the first 8 will be used. For example:

53729.406185

would be stored in the computer as:

53729.406
(8 figures maximum)

Numbers may contain commas in the integer portion, that is, the part before the decimal point. For example:

1,397
4,609,007.2

Commas within numbers should not have spaces around them or they will be taken as separators for items in a list. For example, the following is a list of three numbers:

1,379, 4,297,683, 5,500

Large or small numbers may be written using the short-hand "E-notation". A real number is immediately followed by the letter E and then immediately followed by a positive or negative integer. The integer indicates the power of ten that is to be multiplied by the real number. For example:

9.7E7 is 97000000.
-4.E-5 is -.00004

The "E" stands for "exponent". CONVERSE will print very large and very small real numbers using the E-notation.

4. QUOTATIONS

Quoted words, phrases, and sentences must always be enclosed by quotation marks. Nestings, that is, quotations within quotations, are not allowed. Some examples of quotations are:

WHAT DOES THE WORD "QUOTE" MEAN?

HE SAID, "ON TOP", DIDN'T HE?

"["DOWNEY" IS A CITY.]" IS A SENTENCE.

The last example above is not a nesting since the quotation marks within the square brackets are simply part of the spelling of a complex word.

5. ENTRY CONTROL

A carriage return (*cr*) enters a typed line for processing. If the sentence is not complete (see END OF A SENTENCE), the user will receive another line prompt on the terminal to continue the sentence. The current line being typed may be cancelled by pressing the break (*brk*) key. Characters may be individually deleted by using the backspace (*bs*) or the backslash (\). Only one sentence at a time may be entered. If a hyphen is encountered as the last character in the line, the current syntactic unit being formed will be continued on the next line. In this way, words, numbers, etc. may be broken across two lines.

6. END OF SENTENCE

A sentence will end when one of the following conditions occurs:

- (1) When a ?, !, or ← appears outside square brackets or quotation marks.
- (2) When a period, followed by zero or more blanks, followed by a carriage return, appears outside square brackets or quotation marks.

- (3) When a ?, !, ← , or period, followed immediately by an ending quotation mark, followed by zero or more blanks, followed by a carriage return, appears outside square brackets.

As a result of (2) above, a declarative sentence cannot end in an abbreviation unless a second period is used to end the sentence. For example:

I GO TO U.C.L.A..

As stated in (3) above, a sentence that ends in a quotation may have its terminator appear inside the ending quotation mark if desired. In this case, the terminator is assumed to end the sentence inside the quotation marks as well as the entire sentence. For example:

DID HE SAY, "HOW ARE YOU?"

If different terminators are required, they both must be specified:

YES, HE SAID, "HOW ARE YOU?"!

A sentence that ends in a quoted word should not have its terminator appear inside the marks or the word will be taken as a sentence.

IT'S A "UNICORN".

A sentence cannot be ended if an opening square bracket, parenthesis, or quotation mark is not closed. Parentheses are the only characters that may be nested and all opening parentheses must be closed before the sentence can terminate.

The backarrow (←) will end a sentence, but will not be included as part of it. In this way, partial sentences or sentences without terminators may be entered.

7. DATES AND TIMES

A date is written with slashes between the numbers for the month, day and year.
For example:

12/25/72
4/1/1862

Since the slash is also used as the division symbol (see ARITHMETIC), a date must have all of the following characteristics or it will be taken as two divisions:

- (1) The month must be an integer between 1 and 12 inclusive.
- (2) The day must be an integer between 1 and 31 inclusive.
- (3) The year must be an integer.

A time is written with the colon. For example:

12:42
23:00
5:03:59

The hours must be an integer between 0 and 23 inclusive. The minutes and optional seconds must be integers between 0 and 59 inclusive.

8. ARITHMETIC

The following arithmetic operations are available for the formation of arithmetic expressions:

- + addition
- subtraction
- * multiplication
- / division
- \dagger exponentiation
- % percent

Parentheses may be used for grouping sub-expressions and for altering the precedence of operators (see PRECEDENCE). Words, as well as explicit numbers, may appear in arithmetic expressions provided they have a numeric meaning. Care should be taken when using subtraction and division since the symbols for these operations have other meanings depending on the context. (For other uses of the minus, see WORDS, NUMBERS, and MINUS VS. HYPHEN. For other uses of the slash, see DATES AND TIMES.)

9. PRECEDENCE

Every arithmetic operator has a precedence that determines which one is to be performed first in an expression in which the order is ambiguous. For example, consider the following:

$7+2*3$

Is this expression to be interpreted (a) 7 plus 2 is 9, times 3 is 27, or (b) 2 times 3 is 6, plus 7 is 13? The list below shows the precedence for the arithmetic operators. The highest precedence is listed first.

- (1) \uparrow
- (2) % (left only)
- (3) * /
- (4) + -

Since * has a higher precedence than +, the answer to the above expression example is 13. If two or more operators of the same precedence are in conflict, combination is done from left-to-right. Precedence may be changed by using parentheses to form sub-expressions. For example,

$(7+2)*3$

would be 27.

Certain precedences apply to the various brackets as well as to arithmetic operators. Once inside square brackets, no characters have meaning until the closing square bracket is encountered. Consequently, there is no meaning to square brackets appearing inside themselves. The opening and closing parentheses count is kept separate when inside or outside a quotation. Thus, it is not possible to close an opening parenthesis from inside a quotation if it was opened from outside, and vice versa.

10. MINUS VS. HYPHEN

The following chart shows the distinction between the minus and the hyphen. A determination is made based upon (1) whether a word or a number appears before or after the "-" symbol (which is reflected in the columns of the chart) and (2) whether or not spaces surround it (which is reflected in the rows).

—	X & Y words	X word Y number	X number Y word	X & Y numbers
X-Y	X-Y Hyphenated word	X-Y Hyphenated word	X-Y Hyphenated word	X - Y subtraction
X- Y	X- Y post-hyphen	X- Y post-hyphen	X - Y subtraction	X - Y subtraction
X -Y	X - Y subtraction	X -Y unary minus	X - Y subtraction	X - Y subtraction
X - Y	X - Y subtraction	X - Y subtraction	X - Y subtraction	X - Y subtraction

11. HELP

The hash character (#) is reserved for the user who wants help before the sentence is sent to the CONVERSE parser. The hash may appear either before or after the end of the sentence, or on a line by itself. It will never be included as part of the sentence.

When the hash is encountered, the user is first informed about any open parentheses, open quotations, or open square brackets that are not closed that may exist in the sentence as typed in so far. Second, the sentence is printed since CONVERSE attempts to determine where an error, if any, was made. Third, the user is asked whether he wishes to continue or cancel the sentence. The use of the hash character is the only way a sentence may be cancelled if it has already exceeded one typed line. The break (see ENTRY CONTROL) only cancels the current line being typed.

12. CHARACTER SET

The CONVERSE character set may be divided into five categories as follows:

Set I. This set is composed of the 26 upper case letters and the 10 digits that are used to form words and numbers.

Set II. The characters in this set have one and only one meaning regardless of context. They are individuals and do not combine with other characters to form syntactic units except inside square brackets.

- + addition
- * multiplication
- ↑ exponentiation
- % percent
- ; semicolon
- () parentheses
- " quote marks
- [] square brackets

& ampersand
' apostrophe
! exclamation mark
? question mark
blank space
\$ dollar sign

Set III. This set consists of characters with more than one meaning that is dependent on context.

: colon; time
/ division; dates
, within numbers; separator for items in a list
. period; decimal point; dot for abbreviation;
used to form "..."
- minus; subtraction; hyphen; dash

Set IV. The characters in this set are not presently used but are available for definition.

=
@
<
>

Set V. The last set is composed of control characters that never appear as part of the input sentence.

hash
cr carriage return
bs backspace
brk break
\ backslash
← backarrow

15 March 1972

10

System Development Corporation
TM-4885/000/00

ACKNOWLEDGEMENTS

I would like to acknowledge the following people for
their many suggestions and comments: Charles Kellogg,
Tim Diller, John Burger, and Ken Fogt.